
Quality of Correlating Functions Generated from Commonly Used Basis Sets

T. KOMATSU, T. NORO, and F. SASAKI

Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo, Japan

H. TATEWAKI*

Computation Center, Nagoya City University, Nagoya, Japan

Received 30 June 1995; accepted 17 October 1995

ABSTRACT

Tests have been performed on the quality of correlating functions generated from commonly used Gaussian basis sets, such as the 4-31G and MIDI-4 sets. The atoms tested were carbon, nitrogen, and oxygen. Self-consistent field and configuration interaction (CI) calculations were performed for the ground and lower excited states of neutral atoms as well as for positive and negative ions, using the original sets. Next, after adding (1) one *d*, and (2) two *d* and one *f* primitive Gaussian-type functions (GTFs) to the original sets, the CI calculations were repeated. In order to investigate the quality of the correlating orbitals generated from the GTF sets, parallel calculations to those for the GTF sets were carried out with an extended set of Slater-type functions. It was found that the excitation energies change in a stepwise manner as the basis sets changed from the original sets to the original set +1*d* and the original set +2*d*1*f*. The improvements in excitation energies and ionization energies were almost independent of the original sets and were found to be strongly dependent on the augmented correlation functions. © 1996 by John Wiley & Sons, Inc.

Introduction

Gaussian-type functions (GTFs) have been widely used as basis functions in linear combination of atomic orbital molecular calculations.

*Author to whom all correspondence should be addressed.

Various types of basis sets have been proposed by several authors. Usually, the Gaussian exponents and contraction coefficients in these basis sets are chosen to minimize the self-consistent-field (SCF) energy of the atomic ground states. For larger GTF sets, Almlöf, Taylor, Bauschlicher, and co-workers¹ proposed the basis set taken from the atomic natural orbitals, which provides an excellent descrip-

tion of molecular correlation effects. Dunning² also determined valence *s* and *p* set and *s*, *p*, *d*, *f*, and *g* functions which describe correlation effects by optimization in the atomic correlated calculations. For molecular calculations, however, widely used basis sets are still taken from SCF calculations.

When molecules are considered as an aggregate of atoms, even atoms in the molecular ground state are perturbed by surrounding atoms. The atoms in the molecules might exist in a mixed state of atomic ground and excited states. If the above basis sets are used in molecular calculations, it is implicitly assumed that the basis set can describe several excited and ionized states of atoms, as well as the ground state. When these basis sets were first developed, they were mainly used for molecular SCF calculations. Recently, however, post-SCF calculations, such as Möller-Plesset perturbation and configuration interaction (CI) calculations have become rather common. The basis sets, optimized for the atomic ground states in the SCF scheme, are still used for these post-SCF calculations to assess the electronic correlation effects. At times, polarization functions, such as *d*- and *f*-type functions, are added to the original sets. The exponents of these orbitals are usually chosen to describe the deformation of the valence shell orbitals of the molecular ground state. When developing these basis sets, test calculations were often conducted for molecular systems to demonstrate the performance of the basis set, in most cases using the SCF scheme. It is not known how these basis sets can describe the electron correlations in various molecular states as well as atoms.

In the present paper, test calculations were performed on carbon, nitrogen, and oxygen atoms with commonly used basis sets by SCF-CI calculations to clarify the following properties:

1. The quality of the commonly used basis sets for the SCF wave functions of valence excited states and positive and negative ionized states of atoms.
2. The quality of the correlating functions generated from the original sets for these atomic states.
3. The quality of the correlating functions generated from the augmented polarization functions.

In the next section the methods used in the calculations are explained, together with reference calculations employing extended sets of Slater-type functions (STFs). The results and discussion for the correlation energies and relative energies are given in the third section, followed by concluding remarks.

Methods of the Calculations

Selection was made of the basis sets consisting of 4-31G,^{3,4} MIDI-4,⁵ 6-311G,⁶ and [6111/311] and [42111/311] decontracted from [63/5],⁷ as generated by the method of Tatewaki and Huzinaga.⁸ Either one *d*- (1*d*) or two *d*- and one *f*-type (2*d* + 1*f*) correlating functions were then added to the above basis sets. The prepared GTF sets are listed in Table I. Consideration was then given to the following states of the carbon, nitrogen, and oxygen atoms:

C ⁻	2s ² 2p ³	⁴ S
C	2s ² 2p ²	³ P, ¹ D, ¹ S
C ⁺	2s ² 2p ¹	² P
	2s ¹ 2p ²	⁴ S
N	2s ² 2p ³	⁴ S, ² D, ² P
N ⁺	2s ² 2p ²	³ P, ¹ D, ¹ S
O ⁻	2s ² 2p ⁵	² P
O	2s ² 2p ⁴	³ P, ¹ D, ¹ S
O ⁺	2s ² 2p ³	⁴ S, ² D, ² P

Since nitrogen has no electron affinity, N⁻ was excluded from the present calculations. In all the following calculations, the 1*s* electrons are uncorrelated. The SCF calculations were performed using the GTF sets listed in Table I. Next, SCF plus single and double excitation configuration interaction (SDCI) calculations were conducted with the natural orbital basis (NO-SDCI), which were obtained from SDCI wave functions with SCF orbitals. For the C⁻(⁴S) and O⁻(²P) states, 2p² → 3p² was added as the reference configuration.

Before discussing the results calculated with the GTFs, the SCF-SDCI calculations are explained using the extended STF basis sets which generate the reference energies. The STF sets (18s18p12d10f8g) were prepared. The extensive calculation consists of an SCF calculation and a two-step procedure in CI.

Step 1 of this CI consists of a series of SDCI calculations executed to generate natural orbitals (NOs), where truncation was performed on the NOs with small occupation numbers. For the oxy-

TABLE I.
Names and Numbers Associated with the Primitive GTFs for the Respective CGTF and the Polarization Functions.

No.	Name	Contraction	<i>d</i> Function			<i>f</i> Function		
			C	N	O	C	N	O
1.	4-31G	431 / 31						
2.	MIDI-4	421 / 31						
3.	6-311G	6311 / 311						
4.	[6111 / 311]	6111 / 311						
5.	[42111 / 311]	42111 / 311						
6.	4-31G + <i>d</i>	431 / 31 / 1	0.8	0.8	0.8			
7.	MIDI-4 + <i>d</i>	421 / 31 / 1	0.6	0.964	1.154			
8.	6-311G + <i>d</i>	6311 / 311 / 1	0.626	0.913	1.292			
9.	[6111 / 311] + <i>d</i>	6111 / 311 / 1	0.6	0.864	1.154			
10.	[42111 / 311] + <i>d</i>	42111 / 311 / 1	0.6	0.864	1.154			
11.	6-311G + 2 <i>d</i> + <i>f</i>	6311 / 311 / 11 / 1	0.288	0.412	0.535	0.7	1.0	1.3
			1.335	1.986	2.704			
12.	[6111 / 311] + 2 <i>d</i> + <i>f</i>	6111 / 311 / 11 / 1	0.288	0.412	0.535	0.7	1.0	1.3
			1.335	1.986	2.704			
13.	[42111 / 311] + 2 <i>d</i> + <i>f</i>	42111 / 311 / 11 / 1	0.288	0.412	0.535	0.7	1.0	1.3
			1.335	1.986	2.704			

gen and nitrogen atom, the first SDCI calculation employs the occupied and virtual *s*- and *p*-type SCF orbitals. The second SDCI calculation employs 6–8 *s*- and 5–8 *p*-type NOs from the first SDCI and all the 12 *d*-type STF. The third SDCI calculation uses the *s*- and *p*-type STF NOs from the first calculation, 5–7 *d*-type NOs from the second, and all the *f*-type STF. The same procedure was repeated to produce 4–5 *g*-type NOs. For the carbon atom, the *d*-type NOs were obtained simultaneously with *s*- and *p*-NOs. All the remaining processes in Step 1 are the same as those for the nitrogen and oxygen atoms. The sizes of the NO sets, thus prepared, differ for the respective states.

In Step 2, SDCI calculations were performed using these NO sets. These calculations are termed the extended CIs.

The present STF basis sets are expected to be well saturated, with the resultant NOs close to the accuracy of the NOs in the SDCI limit. Thus, it is worthwhile to construct small-sized basis sets from the NOs with larger occupation numbers, since the sets thus generated can be considered to be the ideal set for a given size of the basis set. In order to compare the energies yielded by the small NO set with those by the same size GTF sets, the CI energies are estimated for the small-sized NO sets using the simple formula described in the Appendix. Hereafter, the NOs of the extended CI are written as STF NOs to identify the function type.

All calculations were performed with the use of the ATOMCI program,⁹ which uses full atomic symmetry.

Results and Discussion

CORRELATION ENERGY

Here, the results obtained by the extended basis sets are discussed. In Table II, a list is constructed of the contributions to the correlation energies from individual STF NOs of the ground states of carbon, nitrogen, and oxygen. It can be seen that each contribution to the correlation energy strongly depends on the principal quantum number *n*, but not on angular momentum *l*. That is, the contribution ordering is

$$3d > 3p > 3s \gg 4f > 4p \sim 4d > 4s \\ \gg 5g \sim 5p > 5f > 5d \dots \quad (1)$$

The first group of *n* = 3 describes important configurations such as $2p^2 \rightarrow 3p^2$, $2s2p \rightarrow 2p3d$, and $2p^2 \rightarrow 3d^2$. To express the correlating orbitals in the first group, at least a split valence (double zeta) basis set and single *d*-type polarization function are required. The second group of *n* = 4 consists of *s*, *p*, *d*, and *f* functions. To cover this group requires a triple zeta plus the two *d* and one *f* function. The third group including *g* function

TABLE II.
Contributions^a from Individual NOs^b
(in millihartrees).

NO	C(³ P)	C-(⁴ S)	N(⁴ S)	O(³ P)	O-(² P)
3s	-13.3	-21.8	-20.7	-25.7	-38.1
3p	-18.1	-49.0	-29.5	-53.4	-129.5
3d	-44.1	-52.9	-62.6	-74.6	-83.1
4s	-1.2	-2.4	-2.2	-2.5	-4.4
4p	-3.5	-6.3	-5.3	-9.7	-24.7
4d	-2.4	-5.2	-4.6	-9.2	-18.2
4f	-5.4	-7.6	-9.7	-18.5	-21.6
5s	-0.2	-0.3	-0.3	-0.4	-0.7
5p	-0.7	-1.8	-1.2	-2.5	-6.8
5d	-0.6	-1.0	-1.0	-1.9	-3.8
5f	-0.7	-1.2	-1.3	-2.9	-4.7
5g	-0.9	-1.2	-1.6	-3.8	-5.2

^aThe energy contributions were estimated by use of Brown's formula.

^bThe NOs were generated from SDCI using the extended STF sets.

still has one-fifth of the contribution of the second group. Table III lists the estimated correlation energies of the ideal small STF NO sets (see the Appendix). The $(3s2p1d)_{\text{STFNO}}$, which includes the first group of correlating orbitals, yields 86, 81, and 74% of the total correlation energies of extended CIs for carbon, nitrogen, and oxygen, respectively. The second group of correlating orbitals results in another 10–18%. On the other hand, if use is not made of the $3d$ orbital, a large portion of the atomic correlation energy is lost. For example, $(3s2p)_{\text{STFNO}}$ yields only 37–46% of the total correlation energy for these atoms. Although we have not shown the correlation contributions, similar results were obtained for the excited and ionized states.

TABLE III.
Estimated Correlation Energies^a (au) of STF NO Sets.

NO Set	C(³ P)	N(⁴ S)	O(³ P)
3s2p	-0.045602 (46.20)	-0.046289 (37.29)	-0.075724 (41.07)
3s2p1d	-0.084974 (86.08)	-0.100064 (80.61)	-0.136688 (74.14)
4s3p2d1f	-0.095289 (96.53)	-0.118011 (95.07)	-0.170294 (92.37)
Full set	-0.098716	-0.12413	-0.184363

^aThe numbers in parentheses indicate the percentage of correlation energies of small NO sets versus those of the full set.

The correlation energies of the ground states of carbon, nitrogen, and oxygen obtained with the GTF sets are shown in Tables IV–VI, respectively. As can be seen, the calculated correlation energies clearly fall into three categories: those of the correlation energies given by the (sp) , $(sp + 1d)$, and $(sp + 2d + 1f)$ sets. Table VII lists the correlation energies obtained by the GTF sets and the small STF NO sets for the ground state of oxygen. The 4-31G and MIDI-4 GTF sets yield correlation energies of -0.0611 and -0.0603 au, respectively, being 0.0146 and 0.0154 au smaller in magnitude than the corresponding $(3s2p)_{\text{STFNO}}$ set. Single d -type polarization increases the correlation energy by 0.055 and 0.061 au for 4-31G and MIDI-4, respectively. The corresponding STF NO $3d$ increases the correlation energy by 0.0610 au. Therefore, the error almost conserved for the MIDI-4 set, namely, the polarization function in MIDI-4, is appropriately expressed by $3d$. On the other hand, the discrepancy for the 4-31G set increased by 0.005 au, where the orbital exponent of the d -type function is 0.8, since the GAUSSIAN program determines the exponent of 0.8 for oxygen as well as for carbon and nitrogen. The replacement of this d function by that of Tatewaki and Huzinaga or that of 6-311G* should result in some improvement. The three kinds of triple zeta GTF sets exhibit common aspects. The (sp) set yields -0.070 au for the correlation energy, which differs by 0.013 au from the value of the STF NO set. The additions of the $1d$ and $2d + 1f$ yield considerable increase of correlation energy, but the difference from the corresponding STF NO set still ranges from 0.012 to 0.014 au. The error of the GTF set relative to the extended calculation mainly results from the poor quality of the (sp) set. The same arguments hold for the excited states and others.

RELATIVE ENERGIES

The excitation energies (EEs), and electron affinities (EAs) for carbon, nitrogen, and oxygen are shown in Figures 1–3, respectively. The ionization energies (IEs) for carbon, nitrogen, and oxygen are also displayed in Figures 4–6, respectively. The IEs and EEs are discussed in this section, and EAs in the section on Negative Ions.

The results given by the respective SCF calculations are quite similar. The present smallest 4-31G and MIDI-4 sets yield relative energies according to those of the extended STF set to within 0.1 eV. Once a given basis set is used, the errors are

TABLE IV.
Total Energies and Correlation Energies for the Ground State of Carbon (in atomic units).

Basis (GTF)	E_{SCF}	E_{CI}	E_{corr}	Percentage ^a
1. 4-31G	-37.635947	-37.674957	-0.039011	39.52
2. MIDI-4	-37.634568	-37.673661	-0.039093	39.60
3. 6-311G	-37.684801	-37.727409	-0.042609	43.16
4. [6111/311]	-37.684508	-37.727018	-0.042510	43.06
5. [42111/311]	-37.684856	-37.727425	-0.042570	43.12
6. 4-31G + <i>d</i>	-37.635947	-37.710059	-0.074112	75.08
7. MIDI-4 + <i>d</i>	-37.634568	-37.712042	-0.077474	78.48
8. 6-311G + <i>d</i>	-37.684801	-37.766183	-0.081382	82.44
9. [6111/311] + <i>d</i>	-37.684508	-37.765692	-0.081184	82.24
10. [42111/311] + <i>d</i>	-37.684856	-37.766082	-0.081226	82.28
11. 6-311G + 2 <i>d</i> + <i>f</i>	-37.684801	-37.775857	-0.091057	92.24
12. [6111/311] + 2 <i>d</i> + <i>f</i>	-37.684508	-37.775249	-0.090742	91.92
13. [42111/311] + 2 <i>d</i> + <i>f</i>	-37.684856	-37.775639	-0.090783	91.96
Dunning ^b	-37.688544	-37.785483	-0.096939	98.20
Extended	-37.688619	-37.787335	-0.098716	100.00

^a $E_{corr}(GTF)/E_{corr}(Extended) \times 100$.^bReference 2.

almost constant for all the states with the sole exception of that for the negative ions, which will be discussed in the next section. Any GTF set which was determined to optimize the ground state SCF energy is capable of representing the excited and ionized states equally as well as the ground state. It should be noted, however, that the SCF calculations at times do not correctly repro-

duce even the order of the energy levels, as shown in Figure 1.

From Figures 1–3, it can be seen that CIs with the original GTF sets and the corresponding STF NO sets give almost the same EEs and IEs. For example, the EE and IE values calculated with the two smallest 4-31G and MIDI-4 sets differ from those of the corresponding STF NO set by 0.2 eV at

TABLE V.
Total Energies and Correlation Energies for the Ground State of Nitrogen (in atomic units).

Basis (GTF)	E_{SCF}	E_{CI}	E_{corr}	Percentage ^a
1. 4-31G	-54.325289	-54.361788	-0.036499	29.40
2. MIDI-4	-54.323482	-54.360000	-0.036518	29.42
3. 6-311G	-54.394737	-54.436689	-0.041952	33.80
4. [6111/311]	-54.394392	-54.435936	-0.041544	33.47
5. [42111/311]	-54.395112	-54.436765	-0.041654	33.56
6. 4-31G + <i>d</i>	-54.325289	-54.414216	-0.088927	71.64
7. MIDI-4 + <i>d</i>	-54.323482	-54.412242	-0.088760	71.51
8. 6-311G + <i>d</i>	-54.394737	-54.489734	-0.094997	76.53
9. [6111/311] + <i>d</i>	-54.394392	-54.489091	-0.094699	76.29
10. [42111/311] + <i>d</i>	-54.395112	-54.489901	-0.094790	76.36
11. 6-311G + 2 <i>d</i> + <i>f</i>	-54.394737	-54.505504	-0.110767	89.23
12. [6111/311] + 2 <i>d</i> + <i>f</i>	-54.394392	-54.504721	-0.110329	88.88
13. [42111/311] + 2 <i>d</i> + <i>f</i>	-54.395112	-54.505528	-0.110417	88.95
Dunning ^b	-54.400771	-54.522023	-0.121252	97.68
Extended	-54.400934	-54.525065	-0.124130	100.00

^a $E_{corr}(GTF)/E_{corr}(Extended) \times 100$.^bReference 2.

TABLE VI.
Total Energies and Correlation Energies for the Ground State of Oxygen (in atomic units).

Basis (GTF)	E_{SCF}	E_{CI}	E_{corr}	Percentage ^a
1. 4-31G	-74.702439	-74.763560	-0.061121	33.15
2. MIDI-4	-74.701604	-74.761899	-0.060295	32.70
3. 6-311G	-74.798619	-74.868224	-0.069605	37.75
4. [6111/311]	-74.798837	-74.868537	-0.069700	37.81
5. [42111/311]	-74.800140	-74.870255	-0.070115	38.03
6. 4-31G + <i>d</i>	-74.702439	-74.818429	-0.115990	62.91
7. MIDI-4 + <i>d</i>	-74.701604	-74.822671	-0.121068	65.67
8. 6-311G + <i>d</i>	-74.798619	-74.930540	-0.131921	71.56
9. [6111/311] + <i>d</i>	-74.798837	-74.930827	-0.131990	71.59
10. [42111/311] + <i>d</i>	-74.800140	-74.932386	-0.132246	71.73
11. 6-311G + 2 <i>d</i> + <i>f</i>	-74.798619	-74.957252	-0.158633	86.04
12. [6111/311] + 2 <i>d</i> + <i>f</i>	-74.798837	-74.957479	-0.158642	86.05
13. [42111/311] + 2 <i>d</i> + <i>f</i>	-74.800140	-74.959046	-0.158906	86.19
Dunning ^b	-74.809068	-74.988205	-0.179137	97.17
Extended	-74.809398	-74.993762	-0.184363	100.00

^a $E_{\text{corr}}(\text{GTF})/E_{\text{corr}}(\text{Extended}) \times 100$.

most. The large portion of errors in the correlation energy of the GTF CIs is canceled when the relative energies are calculated. Recall that for the carbon atom, the SCF calculations failed to give the correct order for the excited states. On the other hand, the CI calculations with the (*sp*) sets gave the correct order. The agreement between the calculated and experimental results,^{10,11} however, is only qualitative.

The addition of the polarization *d* function results in a considerable improvement of the EE and IE values of the CI calculations (see Figs. 1–6). For all the atoms considered, the errors in the calculated EE and IE values relative to experimental results^{10,11} are less than 0.4 eV, with the use of this CI. Further improvement is introduced by the addition of (2*d*1*f*); the errors are less than 0.2 eV. With respect to IEs for a 2*p* ionization of C, values given by the CI with the (*sp*) set are ~ 10.6 eV, while the experimental value is 11.2 eV. For the CI with 3*d* it was found that for C(³P) the contributions of the 2*p*² → 3*d*² and 2*s*2*p* → 2*p*3*d* excitations are large, while that for C⁺(²P), 2*p*² → 3*d*², never appears in two electron excitations. Furthermore, the contribution of 2*s*2*p* → 2*p*3*d* for ²P is much less than that for ³P. The energy lowering contributions from 3*d* are -1.2 and -0.8 eV for ³P and ²P, respectively, yielding an IE of 11.0 eV. Thus, the inclusion of the *d*-type function is essential in reproducing the ionization potential.

The error in the correlation energy of GTF sets relative to that of the extended STF set causes only

a slight effect when the relative energies are calculated. It is felt, however, that the error induces a serious problem when the absolute values of the correlation energies are important, for example, as in geometrical optimizations in molecular calculations. This is discussed in detail, selecting the present largest sets of (*sp* + 2*d* + 1*f*). For the ground states of carbon, nitrogen, and oxygen, the (*sp* + 2*d* + 1*f*) sets yield differences of 0.008, 0.013–0.014, and 0.025–0.026 au relative to the correlation energies given by the extended CI. Borrowing the basis sets of other atoms, the atoms in the molecular environment are expected to increase the atomic correlation energies compared with those of isolated atoms. This step may result in, for example, shorter bond distances and exaggerated binding energies. This is a basis set superposition error^{12–15} (BSSE) in post-SCF calculations.

NEGATIVE IONS

Contrary to the EEs and IEs, the electron affinity (EA) calculated by SCF strongly depends on the basis sets. For carbon, the 4-31G and MIDI-4 sets result in negative EAs of -0.314 and -0.068 eV, respectively. The triple zeta sets yield better EAs of 0.037 ~ 0.252 eV, being considerably smaller than that (0.550 eV) by the extended STF set. This is due to the lack of a diffuse *p*-type function which describes the broader character of the attached electron. To obtain a reliable EA, it is necessary to properly take into account the correlation

TABLE VII.
Comparison of CI Results of STF Calculations for $O(^3P)$ with those of GTF Calculations (in atomic units).

Basis (STF)	E_{CI}	E_{corr}^a	Percentage ^b	Basis (GTF)	E_{SCF}	E_{CI}	E_{corr}	Percentage ^c
1. 3s2p	-74.885122	-0.075724	41.07	1. 4-31G	-74.702439	-74.763560	-0.061121	33.15
2. 4s3p	-74.892367	-0.082969	45.00	2. MIDl-4	-74.701604	-74.761899	-0.060295	32.70
3. 5s3p	-74.892511	-0.083113	45.08	3. 6-311G	-74.798619	-74.868224	-0.069605	37.75
4. 3s2p1d	-74.946087	-0.136688	74.14	4. [6111/311]	-74.798837	-74.868537	-0.069700	37.81
5. 4s3p1d	-74.955835	-0.146436	79.43	5. [42111/311]	-74.800140	-74.870255	-0.070115	38.03
6. 5s3p1d	-74.979853	-0.170455	92.46	6. 4-31G + d	-74.702439	-74.818429	-0.115990	62.91
7. 4s3p2d1f	-74.989871	-0.180473	97.89	7. MIDl-4 + d	-74.701604	-74.822671	-0.121068	65.67
8. 5s3p2d1f	-74.993762	-0.184363	100.00	8. 6-311G + d	-74.798619	-74.930540	-0.131921	71.56
9. Extended				9. [6111/311] + d	-74.798837	-74.930827	-0.131990	71.59
				10. [42111/311] + d	-74.800140	-74.932386	-0.132246	71.73
				11. 6-311G + 2d + f	-74.798619	-74.957252	-0.158633	86.04
				12. [6111/311] + 2d + f	-74.798837	-74.957479	-0.158642	86.05
				13. [42111/311] + 2d + f	-74.800140	-74.959046	-0.158906	86.19

^aThe SCF energy of the STF set is -74.809398 au.

^b $E_{corr}(STF\ NO)/E_{corr}(Extended) * 100$.

^c $E_{corr}(GTF)/E_{corr}(Extended) * 100$.

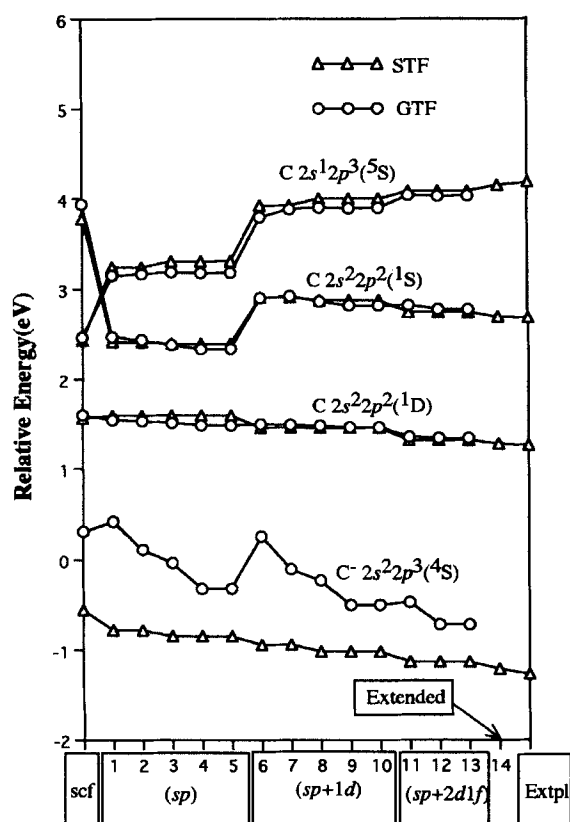


FIGURE 1. Relative energies of the excited states and negatively ionized states of the carbon atom based on the STF (open triangles) and GTF (open circles) sets. The experimental values are taken from refs. 10 and 11.

effects. Here, the energy contributions of $3s$, $3p$, and $3d$ are examined for $C(^3P)$ and $C^-(^4S)$ as listed in Table II. The $3d$ STF NO has the largest contributions of -0.044 and -0.053 au for $C(^3P)$ and $C^-(^4S)$, respectively. The second largest contribution is for $3p$, with values of -0.018 au for C and -0.049 au for C^- . The large contribution for C^- arises from the in-out correlation of $2p^2 \rightarrow 3p^2$. This $3p$ NO also has a diffuse character as does the $2p$ in C^- . The addition of one diffuse p function (0.05) to MIDI-4 lowers the SCF and CI energies of C^- by 0.563 and 0.882 eV, respectively. The EA values in the CI change from -0.116 to 0.729 eV, being within 0.041 eV of the EA by the $(3s2p)_{STFNO}$. This means that the split valence GTF sets plus a diffuse p have an equal quality as the corresponding STF NO set for the negative ion. For the triple zeta set 6-311G, for example, addition of the same diffuse p also yields a good EA of 0.769 eV, within 0.071 eV of the value by the $(4s3p)_{STFNO}$, and differing by 0.5 eV from the observed value of

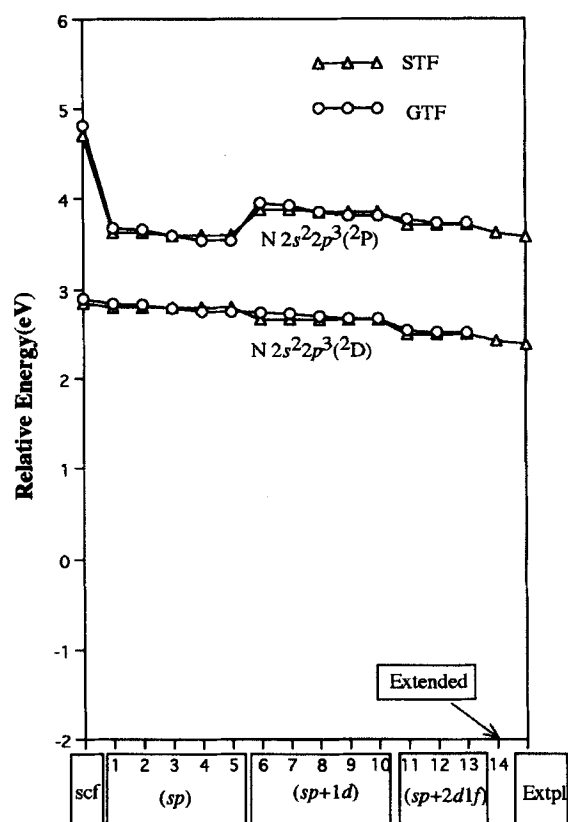


FIGURE 2. Relative energies of the excited states of the nitrogen atom based on the STF (open triangles) and GTF (open circles) sets. The experimental values are taken from refs. 10 and 11.

1.268 eV. The large portion of the remaining error in EA value will be recovered by the addition of polarization functions to any (sp) sets with a diffuse $2p$ GTF. The results for the oxygen atom are similar.

Concluding Remarks

An investigation was conducted of the quality of correlating functions generated from commonly used GTF sets, which were determined to minimize the SCF energy of the atomic ground state. The sets tested were the 4-31G, MIDI-4, 6-311G, [6111/311], and [42111/311]. Furthermore, the original (sp) sets were augmented by one d - and two d - and f -type GTFs. For investigative purposes, the ground, excited, positively, and negatively ionized states of the carbon, nitrogen, and oxygen atoms were chosen. The SDCl calculations from the SCF configurations were performed with

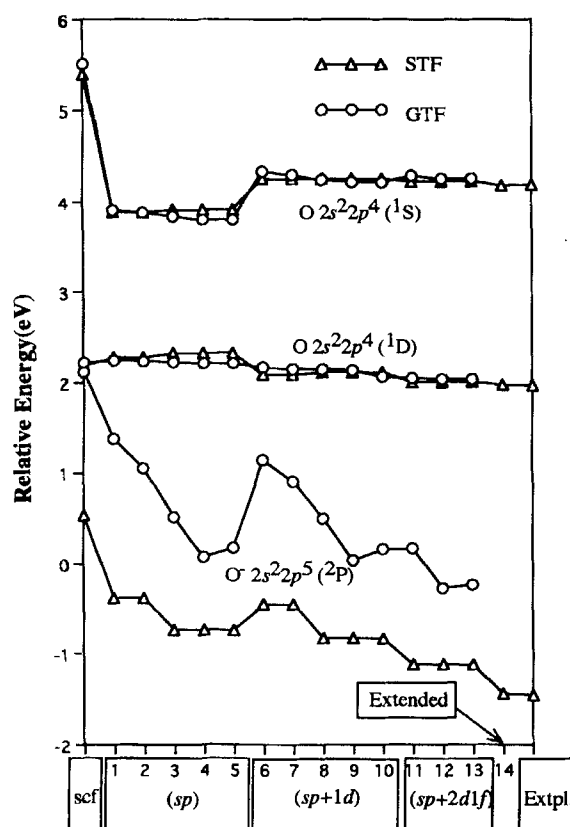


FIGURE 3. Relative energies of the excited states and negatively ionized states of the oxygen atom based on the STF (open triangles) and GTF (open circles) sets. The experimental values are taken from refs. 10 and 11.

these sets. Also SCF-SDCI calculations were conducted using an extended STF basis set and the resultant small sets of STF NOs as references.

The SCF calculations with all GTF basis sets resulted in almost the same excitation energies. It was concluded that the basis sets were quite accurate in the domain of the SCF, but differences from experimental value were considerable. It was found that CI calculations with correlating functions generated from all (*sp*) basis sets yielded nearly the same excitation energies and ionization energies. The errors relative to experimental results were determined as around 1.0 eV for all atoms. Molecular CI calculations with (*sp*) basis sets may produce this order of error in valence-type excitations. The CI calculations with all (*sp* + 1*d*) basis sets resulted in an error of ca. 0.3 eV for the atom under consideration. The same accuracy can be expected in molecular calculations. The (*sp* + 2*d* + 1*f*) calculations were performed for the 6-311G,

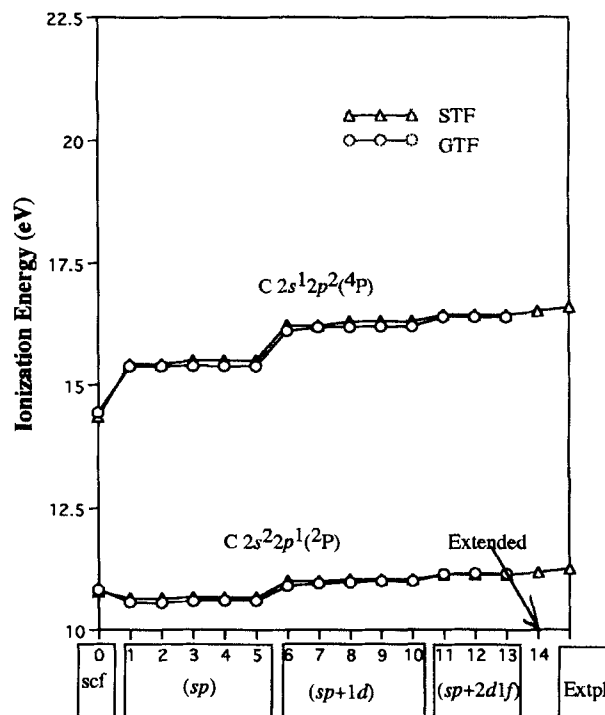


FIGURE 4. Ionization energies of the carbon atom based on the STF (open triangles) and GTF (open circles) sets. The experimental values are taken from ref. 10.

[6111/311], and [42111/311] sets. The errors for the excitation energies and ionization energies were found to be within 0.2 eV, and considered satisfactorily small. The (*sp* + 1*d*) and (*sp* + 2*d* + 1*f*) sets are, therefore, recommended in the calculation of molecular valence excited states.

We also found that all the GTF sets gave almost the same results as the corresponding ideal STF set. Although we have not performed the GTF calculations with the *g*-type GTFs, we can expect that the effects of the *g* functions in the GTF calculations are the same as those in the corresponding ideal STF, where the triple zeta (*sp*) basis such as [6111/311] and [42111/311] are at least required [see relation (1) and Table II].

Recall that for the negative ion, all the GTF calculations under consideration failed to obtain proper electron affinities, since they did not include diffuse GTFs. When considering the calculation of molecular states, including ionic states as well as negative ions, the inclusion of diffuse orbitals is indispensable.

The data of correlation energies and relative energies of various states are available from the authors.

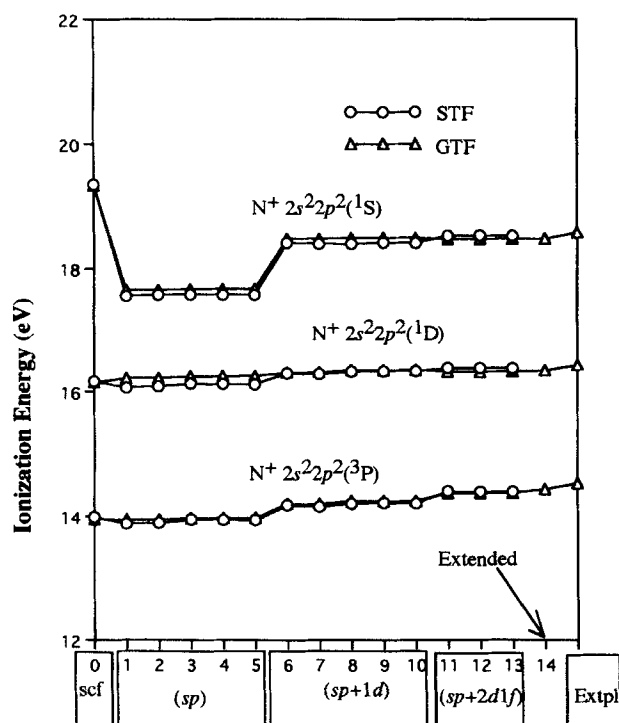


FIGURE 5. Ionization energies of the nitrogen atom based on the STF (open triangles) and GTF (open circles) sets. The experimental values are taken from ref. 10.

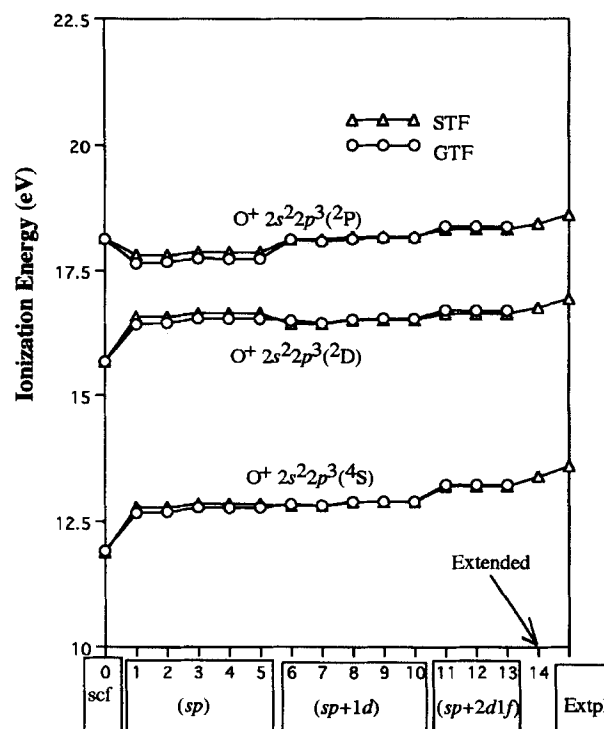


FIGURE 6. Ionization energies of the oxygen atom, based on the STF (open triangles) and GTF (open circles) sets. The experimental values are taken from ref. 10.

Appendix

ESTIMATION OF ENERGY CONTRIBUTIONS TO THE TOTAL ENERGY

Throughout this study, the contribution of a correlating orbital to the total energy was evaluated as follows.

This study was not concerned with exact values but rather the relative importance of the energy contribution of each orbital. Therefore, the orbital contribution was estimated by simply summing the energy contributions of the CSFs that involved the correlating orbital in question. Each CSF's contribution to the total energy was estimated by using Brown's formula¹⁶:

$$\Delta E_i = \frac{(E - H_{ii})|c_i|^2}{\langle \Psi | \Psi \rangle - |c_i|^2} \quad (\text{A1})$$

where Ψ is a CI wave function,

$$\Psi = \sum_i c_i \Phi_i, \quad (\text{A2})$$

and E is its eigenvalue

$$\sum_j H_{ij} c_j = E c_i \quad (\text{A3})$$

Equation (A1) can be deduced by evaluating the expectation value of the wave function, where the relevant CSF has been deleted from the CI expansion (A2) while the remaining coefficients are unchanged.

Note that the sum of the contributions from all correlating orbitals does not correspond to the total correlation energy, since the effects of some CSFs are counted more than once.

Small NO set energies ($3s2p$, $4s3p$, ..., $5s3p2d1f$, etc.) of the external STF basis are evaluated in a similar manner. Starting from the full CI calculations ($6-8s$, $5-8p$, $5-7d$, $4-6f$, $4-5g$ NOs), the CSF's contributions excluded in the small NO set are subtracted. Table VIII compares the estimated energies to the true truncated CI energies. The differences are, at most, 0.005 au for $4s3p$ and 0.0001 au for $5s3p2f1d$; both values are accurate enough for the present purpose.

TABLE VIII.

Comparison between CI Energy and the Estimated Energy for the Various States of the Carbon, Nitrogen, and Oxygen Atoms (in atomic units).

State	4s3p		5s3p2d1f	
	CI Energy	Estimated Energy	CI Energy	Estimated Energy
C(³ P)	-37.733709	-37.737226	-37.783987	-37.783996
C(¹ D)	-37.679239	-37.678651	-37.735459	-37.735556
C(¹ S)	-37.648510	-37.649147	-37.683080	-37.683121
C(⁵ S)	-37.615391	-37.615571	-37.633931	-37.633937
C ⁺ (² P)	-37.344980	-37.345478	-37.374603	-37.374634
C ⁺ (⁴ P)	-37.167414	-37.167526	-37.179771	-37.179771
C ⁻ (⁴ S)	-37.764920	-37.768112	-37.825539	-37.825553
N(⁴ S)	-54.446837	-54.451569	-54.519047	-54.519101
N(² D)	-54.345384	-54.348692	-54.427219	-54.427347
N(² P)	-54.316937	-54.319470	-54.382719	-54.382765
N ⁺ (³ P)	-53.934452	-53.938028	-53.990626	-53.990635
N ⁺ (¹ D)	-53.854725	-53.854449	-53.854817	-53.918970
N ⁺ (¹ S)	-53.801844	-53.802473	-53.839722	-53.839755
O(³ P)	-74.884561	-74.892367	-74.979786	-74.979853
O(¹ D)	-74.803657	-74.806975	-74.906023	-74.906026
O(¹ S)	-74.745579	-74.748496	-74.824759	-74.824610
O ⁺ (⁴ S)	-74.415248	-74.419917	-74.494370	-74.494396
O ⁺ (² D)	-74.276977	-74.280285	-74.368025	-74.368115
O ⁺ (² P)	-74.233067	-74.235327	-74.305515	-74.305533
O ⁻ (² P)	-74.920406	-74.919425	-75.024400	-75.021042

References

1. J. Almlöf and P. R. Taylor, *J. Chem. Phys.*, **86**, 4070 (1987); C. W. Bauschlicher, Jr. and P. Taylor, *J. Chem. Phys.*, **86**, 858, 1420, 2844, 5600 (1987); C. W. Bauschlicher, Jr. and S. R. Langhoff, *J. Chem. Phys.*, **86**, 5595 (1987); C. W. Bauschlicher, Jr., S. R. Langhoff, P. R. Taylor, and R. L. Jaffe, *J. Chem. Phys.*, **88**, 1743 (1988).
2. T. H. Dunning, Jr., *J. Chem. Phys.*, **90**, 1009 (1989).
3. R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.*, **54**, 724 (1971).
4. W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, **56**, 4223 (1971).
5. H. Tatewaki and S. Huzinaga, *J. Comput. Chem.*, **1**, 205 (1980).
6. R. Krishnan, B. S. Binkley, R. Seeger, and J. A. Pople, *J. Chem. Phys.*, **72**, 650 (1980).
7. S. Huzinaga, J. Andzelm, M. Klobukowski, E. Radzio-Andzelm, Y. Sakai, and H. Tatewaki, *Gaussian Basis Sets for Molecular Calculations*, Elsevier, New York, 1984.
8. H. Tatewaki and S. Huzinaga, *J. Chem. Phys.*, **71**, 4339 (1979).
9. F. Sasaki, M. Sekiya, T. Noro, K. Ohtsuki, and Y. Osanai, *MOTEC-91*, ESCOM Science Publishers B.V., Leiden, Netherlands, 1991, p. 115.
10. S. Bashkin and J. D. Stoner, Jr., *Atomic Energy Levels and Grotian Diagrams 1*, North-Holland, Amsterdam, 1975.
11. H. Hotop and W. C. Lineberger, *J. Phys. Chem. Ref. Data*, **4**, 539 (1975).
12. S. F. Boys and F. Bernardy, *Mol. Phys.*, **19**, 553 (1970).
13. A. Meunier, B. Levy, and G. Berthier, *Theor. Chim. Acta (Berlin)*, **29**, 49 (1973).
14. A. Johansson, P. Kollman, and S. Rothenberg, *Theor. Chim. Acta (Berlin)*, **29**, 167 (1973).
15. M. Urban and P. Hobza, *Theor. Chim. Acta (Berlin)*, **36**, 215 (1975).
16. R. E. Brown, Ph.D. thesis, Indiana University, Bloomington, Indiana, 1967.